

SEQUENCE LISTING

SEQ ID NO:1 Synthetic Sso7d gene

GCAACCGTAAAGTTCAAGTACAAAGGCGAAGAAAAAGAGGTAGACATCTCCAA
5 GATCAAGAAAGTATGGCGTGTGGGCAAGATGATCTCCTTCACCTACGACGAGGG
CGGTGGCAAGACCGGCCGTGGTGCAGTAAGCGAAAAGGACGCGCCGAAGGAGC
TGCTGCAGATGCTGGAGAAG
CAGAAAAAG

10 SEQ ID NO:2 The amino acid sequence of Sso7d.

ATVKFKYKGEEKEVDISKIKKVWRVGKMISFTYDEGGGKTGRGAVSEKDAPKELLQ
MLEKQKK

SEQ ID NO:3 The DNA sequence encoding the Sso7d-ΔTaq fusion protein

15 ATGATTACGAATTCGAGCGCAACCGTAAAGTTCAAGTACAAAGGCGAAGAAAAA
GAGGTAGACATCTCCAAGATCAAGAAAGTATGGCGTGTGGGCAAGATGATCTCC
TTCACCTACGACGAGGGCGGTGGCAAGACCGGCCGTGGTGCAGTAAGCGAAAAG
GACGCGCCGAAGGAGCTGCTGCAGATGCTGGAGAAGCAGAAAAAGGGCGGCGG
TGTCAGTAGTCCCAAGGCCTGGAGGAGGCCCCCTGGCCCCCGCCGGAAGGGGGCC
20 TTCGTGGGCTTTGTGCTTTCCCGCAAGGAGCCCATGTGGGCCGATCTTCTGGCCCT
GGCCGCGCCAGGGGGGGCCGGGTCCACCGGGCCCCCGAGCCTTATAAAGCCCT
CAGGGACCTGAAGGAGGCGCGGGGGCTTCTCGCCAAAGACCTGAGCGTTCTGGC
CCTGAGGGAAGGCCTTGGCCTCCCGCCCGGCGACGACCCCATGCTCCTCGCCTAC
CTCCTGGACCCTTCCAACACCACCCCGAGGGGGTGGCCCGGCGCTACGGCGGG
25 GAGTGGACGGAGGAGGCGGGGGAGCGGGCCGCCCTTCCGAGAGGCTCTTCGCC
AACCTGTGGGGGAGGCTTGAGGGGGAGGAGAGGCTCCTTTGGCTTTACCGGGAG
GTGGAGAGGCCCCCTTCCGCTGTCCTGGCCACATGGAGGCCACGGGGGTGCGC
CTGGACGTGGCCTATCTCAGGGCCTTGTCCTGGAGGTGGCCGAGGAGATCGCCC
GCCTCGAGGCCGAGGTCTTCCGCCTGGCCGGCCACCCCTTCAACCTCAACTCCCG
30 GGACCAGCTGGAAAGGGTCCTCTTTGACGAGCTAGGGCTTCCCGCCATCGGCAA
GACGGAGAAGACCGGCAAGCGCTCCACCAGCGCCGCGTCCTGGAGGCCCTCCG
CGAGGCCACCCCATCGTGGAGAAGATCCTGCAGTACCGGGAGCTACCAAGCT
GAAGAGCACCTACATTGACCCCTTGCCGGACCTCATCCACCCAGGACGGGGCCG
CCTCCACACCCGCTTCAACCAGACGGCCACGGCCACGGGCAGGCTAAGTAGCTC

CGATCCCAACCTCCAGAACATCCCCGTCCGCACCCCGCTTGGGCAGAGGATCCGC
CGGGCCTTCATCGCCGAGGAGGGGTGGCTATTGGTGGCCCTGGACTATAGCCAG
ATAGAGCTCAGGGTGCTGGCCACCTCTCCGGCGACGAGAACCTGATCCGGGTCT
TCCAGGAGGGGCGGGACATCCACACGGAGACCGCCAGCTGGATGTTTCGGCGTCC
5 CCCGGGAGGCCGTGGACCCCCTGATGCGCCGGGCGGCCAAGACCATCAACTTCG
GGGTCCTCTACGGCATGTCGGCCACCGCCTCTCCCAGGAGCTAGCCATCCCTTA
CGAGGAGGCCCAGGCCTTCATTGAGCGCTACTTTCAGAGCTTCCCCAAGGTGCGG
GCCTGGATTGAGAAGACCCTGGAGGAGGGCAGGAGGCGGGGGTACGTGGAGAC
CCTCTTCGGCCGCCGCCGCTACGTGCCAGACCTAGAGGCCCGGGTGAAGAGCGT
10 GCGGGAGGCGGCCGAGCGCATGGCCTTCAACATGCCCCGTCCAGGGCACCGCCGC
CGACCTCATGAAGCTGGCTATGGTGAAGCTCTTCCCCAGGCTGGAGGAAATGGG
GGCCAGGATGCTCCTTCAGGTCCACGACGAGCTGGTCCTCGAGGCCCAAAAGA
GAGGGCGGAGGCCGTGGCCCGGCTGGCCAAGGAGGTCATGGAGGGGGTGTATCC
CCTGGCCGTGCCCCTGGAGGTGGAGGTGGGGATAGGGGAGGACTGGCTCTCCGC
15 CAAGGAGGGCATTGATGGCCGCGGCGGAGGCGGGCATCATCATCATCATTA
A

SEQ ID NO:4 The amino acid sequence of Sso7d-ΔTaq fusion protein

MITNSSATVKFKYKGEEKEVDISKIKKVWRVGKMISFTYDEGGGKTGRGAVSEKDA
20 PKELLQMLEKQKKGGGVTSKALEEAPWPPPEGAFVGFVLSRKEPMWADLLALAA
ARGGRVHRAPEPYKALRDLKEARGLLAKDLSVLALREGLGLPPGDDPMLLAYLLDP
SNTTPEGVARRYGGEWTEEAGERAAALSERLFANLWGRLEGEERLLWLYREVERPLS
AVLAHMEATGVRLDVAYLRALSLEVAEEIARLEAEVFRLAGHPFNLNSRDQLERVLF
DELGLPAIGKTEKTGKRSTSAAVLEALREAHPIVEKILQYRELTKLKSTYIDPLPLIH
25 PRTGRLHTRFNQTATATGRLSSSDPNLQNIPVRTPLGQRIRRAFIAEEGWLLVALDYS
QIELRVLAHLSGDENLIRVFQEGRDIHTETASWMFGVPREAVDPLMRRAAKTINFGV
LYGMSAHRLSQELAIPIYEEAQAFIERYFQSFQKVRWIEKTLEEGRRRGYVETLFGRR
RYVPDLEARVKSREAAERMAFNMPVQGTAAADLMKLAMVKLFPRLEEMGARMMLL
QVHDELVLEAPKERAEEAVARLAKEVMGVYPLAVPLEVEVGIGEDWLSAKEGIDGR
30 GGGGHHHHHH

SEQ ID NO:5 The DNA sequence encoding the Sso7d-Taq fusion protein

ATGATTACGAATTCGAGCGCAACCGTAAAGTTCAAGTACAAAGGCGAAGAAAAA

GAGGTAGACATCTCCAAGATCAAGAAAGTATGGCGTGTGGGCAAGATGATCTCC
TTCACCTACGACGAGGGCGGTGGCAAGACCGGCCGTGGTGCAGTAAGCGAAAAG
GACGCGCCGAAGGAGCTGCTGCAGATGCTGGAGAAGCAGAAAAAGGGCGGCGG
TGTCCTAGTGGGATGCTGCCCCCTCTTTGAGCCCAAGGGCCGGGTCCTCCTGGTG
5 GACGGCCACCACCTGGCCTACCGCACCTTCCACGCCCTGAAGGGCCTCACCACCA
GCCGGGGGGAGCCGGTGCAGGCGGTCTACGGCTTCGCCAAGAGCCTCCTCAAGG
CCCTCAAGGAGGACGGGGACGCGGTGATCGTGGTCTTTGACGCCAAGGCCCCCT
CCTTCCGCCACGAGGCCTACGGGGGGGTACAAGGCGGGCCGGGCCCCCACGCCAG
AGGACTTTCCCCGGCAACTCGCCCTCATCAAGGAGCTGGTGGACCTCCTGGGGCT
10 GCGCGCCTCGAGGTCCCCGGGCTACGAGGCGGACGACGTCCTGGCCAGCCTGGC
CAAGAAGGCGGAAAAGGAGGGCTACGAGGTCCGCATCCTCACCGCCGACAAAG
ACCTTTACCAGCTCCTTTCCGACCGCATCCACGTCCTCCACCCCGAGGGGTACCT
CATCACCCCGGCCTGGCTTTGGGAAAAGTACGGCCTGAGGCCCCGACCAGTGGGC
CGACTACCGGGGCCCTGACCGGGGACGAGTCCGACAACCTTCCCGGGGTCAAGGG
15 CATCGGGGAGAAGACGGCGAGGAAGCTTCTGGAGGAGTGGGGGAGCCTGGAAG
CCCTCCTCAAGAACCTGGACCGGCTGAAGCCCGCCATCCGGGAGAAGATCCTGG
CCCACATGGACGATCTGAAGCTCTCCTGGGACCTGGCCAAGGTGCGCACCGACCT
GCCCCTGGAGGTGGACTTCGCCAAAAGGCGGGAGCCCGACCGGGAGAGGCTTAG
GGCCTTTCTGGAGAGGCTTGAGTTTGGCAGCCTCCTCCACGAGTTCGGCCTTCTG
20 GAAAGCCCCAAGGCCTGGAGGAGGCCCCCTGGCCCCCGCCGGAAGGGGCCTTC
GTGGGCTTTGTGCTTTCCCGCAAGGAGCCCATGTGGGCGGATCTTCTGGCCCTGG
CCGCCGCCAGGGGGGGCCGGGTCCACCGGGCCCCCGAGCCTTATAAAGCCCTCA
GGGACCTGAAGGAGGCGCGGGGGCTTCTCGCCAAAGACCTGAGCGTTCTGGCCC
TGAGGGAAGGCCTTGGCCTCCCGCCCGGCGACGACCCCATGCTCCTCGCCTACCT
25 CCTGGACCCTTCCAACACCACCCCCGAGGGGGTGGCCCCGGCGCTACGGCGGGGA
GTGGACGGAGGAGGCGGGGGAGCGGGCCGCCCTTCCGAGAGGCTCTTCGCCAA
CCTGTGGGGGAGGCTTGAGGGGGAGGAGAGGCTCCTTTGGCTTTACCGGGAGGT
GGAGAGGCCCCCTTCCGCTGTCCTGGCCACATGGAGGCCACGGGGGTGCGCCT
GGACGTGGCCTATCTCAGGGCCTTGTCCCTGGAGGTGGCCGAGGAGATCGCCCG
30 CCTCGAGGCCGAGGTCTTCCGCTGGCCGGCCACCCCTTCAACCTCAACTCCCGG
GACCAGCTGGAAAGGGTCCTCTTTGACGAGCTAGGGCTTCCCGCCATCGGCAAG
ACGGAGAAGACCGGCAAGCGCTCCACCAGCGCCGCGCTCCTGGAGGCCCTCCGC
GAGGCCCACCCCATCGTGGAGAAGATCCTGCAGTACCGGGAGCTCACCAAGCTG
AAGAGCACCTACATTGACCCCTTGCCGGACCTCATCCACCCAGGACGGGCGGCC

TCCACACCCGCTTCAACCAGACGGCCACGGCCACGGGCAGGCTAAGTAGCTCCG
 ATCCCAACCTCCAGAACATCCCCGTCCGCACCCCGCTTGGGCAGAGGATCCGCCC
 GGCCTTCATCGCCGAGGAGGGGTGGCTATTGGTGGCCCTGGACTATAGCCAGAT
 AGAGCTCAGGGTGCTGGCCACCTCTCCGGCGACGAGAACCTGATCCGGGTCTTC
 5 CAGGAGGGGCGGGACATCCACACGGAGACCGCCAGCTGGATGTTTCGGCGTCCCC
 CGGGAGGGCCGTGGACCCCTGATGCGCCGGGCGGGCAAGACCATCAACTTCGGG
 GTCCTCTACGGCATGTCGGCCACCGCCTCTCCCAGGAGCTAGCCATCCCTTACG
 AGGAGGCCCAGGCCTTCATTGAGCGCTACTTTCAGAGCTTCCCCAAGGTGCGGGG
 CTGGATTGAGAAGACCCTGGAGGAGGGCAGGAGGCGGGGGTACGTGGAGACCC
 10 TCTTCGGCCGCCGCGCTACGTGCCAGACCTAGAGGCCCGGGTGAAGAGCGTGC
 GGGAGGCGGCCGAGCGCATGGCCTTCAACATGCCCGTCCAGGGCACCGCCGCCG
 ACCTCATGAAGCTGGCTATGGTGAAGCTCTTCCCCAGGCTGGAGGAAATGGGGG
 CCAGGATGCTCCTTCAGGTCCACGACGAGCTGGTCCTCGAGGCCCAAAAGAGA
 GGGCGGAGGCCGTGGCCCGGCTGGCCAAGGAGGTCATGGAGGGGGTGTATCCCC
 15 TGGCCGTGCCCCTGGAGGTGGAGGTGGGGATAGGGGAGGACTGGCTCTCCGCCA
 AGGAGGGCATTGATGGCCGCGGCGGAGGCGGGCATCATCATCATCATTA

SEQ ID NO:6 The amino acid sequence of Sso7d-Taq fusion protein.

MITNSSATVKFKYKGEEKEVDISKIKKVWRVGKMISFTYDEGGGKTGRGAVSEKDA
 20 PKELLQMLEKQKKGGGVTSFMLPLFEPKGRVLLVDGHHLAYRTFHALKGLTTSRGE
 PVQAVYGFASLLKALKEDGDAVIVVFDKAPSFRHEAYGGYKAGRAPTPEDFPRQ
 LALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQLLSDR
 IHVLHPEGYLITPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKGIGECTARKLL
 EEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDFAKRREP
 25 DRERLRAFLEFGLLESPLKALEEAPWPPPEGAFVGVFLSRKEPMWADL
 LALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLSVLALREGLGLPPGDDPMLLA
 YLLDPSNTTPEGVARRYGGEWTEEAGERALSERLFANLWGRLEGEERLLWLYREV
 ERPLSAVLAHMEATGVRLDVAYLRALSLEVAEEIARLEAEVFRLAGHPFNLNSRDQL
 ERVLFDELGLPAIGKTEKTGKRSTSAVLEALREAHPIVEKILQYRELTKLKSTYIDPL
 30 PDLIHPRTGRLHTRFNQTATATGRLSSSDPNLQNPVRTPLGQRIRRAFAIEEGWLLVA
 LDYSQIELRVLAHLSGDENLIRVFQEGRDIHTETASWMFGVPREAVDPLMRRAAKTI
 NFGVLYGMSAHRLSQELAIPIYEEAQAFIERYFQSFPKVRAWIEKTLEEGRRRGYVETL
 FGRRRYVPDLEARVKSVREAAERMAFNMPVQGTAAADLMKLAMVKLFPRL EEMGA

RMLLQVHDELVLEAPKERAEAVARLAKEVMEGVYPLAVPLEVEVGIGEDWLSAKE
GIDGRGGGGHHHHHH

SEQ ID NO:7 The DNA sequence encoding the Pfu-Sso7d fusion protein

5 ATGATTTTAGATGTGGATTACATAACTGAAGAAGGAAAACCTGTTATTAGGCTAT
TCAAAAAAGAGAACGGAAAATTTAAGATAGAGCATGATAGAACTTTTAGACCAT
ACATTTACGCTCTTCTCAGGGATGATTCAAAGATTGAAGAAGTTAAGAAAATAAC
GGGGGAAAGGCATGGAAAGATTGTGAGAATTGTTGATGTAGAGAAGGTTGAGAA
AAAGTTTCTCGGCAAGCCTATTACCGTGTGGAAACTTTATTTGGAACATCCCCAA
10 GATGTTCCCACTATTAGAGAAAAAGTTAGAGAACATCCAGCAGTTGTGGACATCT
TCGAATACGATATTCCATTTGCAAAGAGATACCTCATCGACAAAGGCCTAATACC
AATGGAGGGGGAAGAAGAGCTAAAGATTCTTGCCTTCGATATAGAAACCCTCTA
TCACGAAGGAGAAGAGTTTGGAAGAGGCCCAATTATAATGATTAGTTATGCAGA
TGAAAATGAAGCAAAGGTGATTACTTGGAACCATAGATCTTCCATACGTTGA
15 GGTGTATCAAGCGAGAGAGAGATGATAAAGAGATTTCTCAGGATTATCAGGGA
GAAGGATCCTGACATTATAGTTACTTATAATGGAGACTCATTGACTTCCCATAT
TTAGCGAAAAGGGCAGAAAACTTGGGATTAAATTAACCATTGGAAGAGATGGA
AGCGAGCCCAAGATGCAGAGAATAGGCGATATGACGGCTGTAGAAGTCAAGGG
AAGAATACATTTGACTTGTATCATGTAATAACAAGGACAATAAATCTCCCAACA
20 TACACACTAGAGGCTGTATATGAAGCAATTTTGGAAAGCCAAAGGAGAAGGTA
TACGCCGACGAGATAGCAAAAGCCTGGGAAAGTGGAGAGAACCTTGAGAGAGTT
GCCAAATACTCGATGGAAGATGCAAAGGCAACTTATGAACTCGGGAAAGAATTC
CTTCCAATGGAAATTCAGCTTTCAAGATTAGTTGGACAACCTTTATGGGATGTTT
CAAGGTCAAGCACAGGGAACCTTGTAGAGTGGTTCTTACTTAGGAAAGCCTACG
25 AAAGAAACGAAGTAGCTCCAAACAAGCCAAGTGAAGAGGAGTATCAAAGAAGG
CTCAGGGAGAGCTACACAGGTGGATTTCGTTAAAGAGCCAGAAAAGGGGTTGTGG
GAAAACATAGTATACCTAGATTTTAGAGCCCTATATCCCTCGATTATAATTACCC
ACAATGTTTCTCCCGATACTCTAAATCTTGAGGGATGCAAGAACTATGATATCGC
TCCTCAAGTAGGCCACAAGTTCTGCAAGGACATCCCTGGTTTTATACCAAGTCTC
30 TTGGGACATTTGTTAGAGGAAAGACAAAAGATTAAAGACAAAATGAAGGAAACT
CAAGATCCTATAGAAAAAATACTCCTTGACTATAGACAAAAGCGATAAACTC
TTAGCAAATTCTTTCTACGGATATTATGGCTATGCAAAAGCAAGATGGTACTGTA
AGGAGTGTGCTGAGAGCGTTACTGCCTGGGGAAGAAAGTACATCGAGTTAGTAT
GGAAGGAGCTCGAAGAAAAGTTTGGATTAAAGTCCTCTACATTGACACTGATG

GTCTCTATGCAACTATCCCAGGAGGAGAAAGTGAGGAAATAAAGAAAAAGGCTC
TAGAATTTGTAAAATACATAAATTCAAAGCTCCCTGGACTGCTAGAGCTTGAATA
TGAAGGGTTTTATAAGAGGGGATTCTTCGTTACGAAGAAGAGGTATGCAGTAAT
AGATGAAGAAGGAAAAGTCATTACTCGTGGTTTAGAGATAGTTAGGAGAGATTG
5 GAGTGAAATTGCAAAAGAACTCAAGCTAGAGTTTTGGAGACAATACTAAAACA
CGGAGATGTTGAAGAAGCTGTGAGAATAGTAAAAGAAGTAATACAAAAGCTTGC
CAATTATGAAATTCCACCAGAGAAGCTCGCAATATATGAGCAGATAACAAGACC
ATTACATGAGTATAAGGCGATAGGTCCTCACGTAGCTGTTGCAAAGAACTAGCT
GCTAAAGGAGTTAAAATAAAGCCAGGAATGGTAATTGGATACATAGTACTTAGA
10 GGCGATGGTCCAATTAGCAATAGGGCAATTCTAGCTGAGGAATACGATCCCAA
AAGCACAAGTATGACGCAGAATATTACATTGAGAACCAGGTTCTTCCAGCGGTA
CTTAGGATATTGGAGGGATTTGGATACAGAAAGGAAGACCTCAGATACCAAAAG
ACAAGACAAGTCGGCCTAACTTCCTGGCTTAACATTAAAAAATCCGGTACCGGC
GGTGGCGGTGCAACCGTAAAGTTCAAGTACAAAGGCGAAGAAAAAGAGGTAGA
15 CATCTCCAAGATCAAGAAAGTATGGCGTGTGGGCAAGATGATCTCCTTCACCTAC
GACGAGGGCGGTGGCAAGACCGGCCGTGGTGCGGTAAGCGAAAAGGACGCGCC
GAAGGAGCTGCTGCAGATGCTGGAGAAGCAGAAAAAGTGA

SEQ ID NO:8 The amino acid sequence of the Pfu-Sso7d fusion protein

20 MILDVDYITEEGKPVIRLFKKENGKFKIEHRTFRPYIYALLRDDSKIEEVKKITGERH
GKIVRIVDVEKVEKKFLGKPITVWKLYLEHPQDVPTIREKVREHPAVVDIFEYDIPFA
KRYLIDKGLIPMEGEEELKILAFDIETLYHEGEEFGKGPIIMISYADENEAKVITWKNID
LPYVEVVSSEREMIKRFLRIIREKDPDIIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDGS
EPKMQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYEAIFGKPKKEKVYADEI
25 AKAWESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSTGN
LVEWFLLRKAYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLWENIVYLDFR
ALYPSIIITHNVSPDTLNLEGCKNYDIAPQVGHKFCCKDIPGFIPSLLGHLLEERQKIKTK
MKETQDPIEKILLDYRQKAIKLLANSFYGYGYAKARWYCKECAESVTAWGRKYIE
LVWKELEEKFGFKVLYIDTDGLYATIPGGESEIKKKALEFVKYINSKLPGLLELEYE
30 GFYKRGFFVTKKRYAVIDEEGKVITRGGLEIVRRDWSEIAKETQARVLETILKHGDVEE
AVRIVKEVIQKLANYEIPPEKLAIYEQITRPLHEYKAIGPHVAVAKKLAAGVKIKPG
MVIGYIVLRGDGPISNRAILAEYDPKKHKYDAEYYIENQVLPVLRILEGFGYRKED
LRYQKTRQVGLTSWLNIKKSGTGGGGATVKFKYKGEEKEVDISKIKKVWRVVGKMIS
FTYDEGGGKTGRGAVSEKDAPKELLQMLEKQKK

SEQ ID NO:9 The DNA sequence encoding the Sac7d-ΔTaq fusion protein

ATGATTACGAATTCGACGGTGAAGGTAAAGTTCAAGTATAAGGGTGAAGAGAAA
GAAGTAGACACTTCAAAGATAAAGAAGGTTTGGAGAGTAGGCAAAATGGTGTCC
5 TTTACCTATGACGACAATGGTAAGACAGGTAGAGGAGCTGTAAGCGAGAAAGAT
GCTCCAAAAGAATTATTAGACATGTTAGCAAGAGCAGAAAGAGAGAAGAAAGG
CGGCGGTGTCACTAGTCCCAAGGCCCTGGAGGAGGCCCCCTGGCCCCCGCCGGA
AGGGGCCCTTCGTGGGCTTTGTGCTTTCCCGCAAGGAGCCCATGTGGGCCGATCTT
CTGGCCCTGGCCGCCGCCAGGGGGGGCCGGGTCCACCGGGCCCCCGAGCCTTAT
10 AAAGCCCTCAGGGACCTGAAGGAGGCGCGGGGGCTTCTCGCCAAAGACCTGAGC
GTTCTGGCCCTGAGGGAAGGCCTTGGCCTCCCGCCCGGCGACGACCCCATGCTCC
TCGCCTACCTCCTGGACCCCTTCCAACACCACCCCGAGGGGGTGGCCCGGCGCTA
CGGCGGGGAGTGGACGGAGGAGGCGGGGGAGCGGGCCGCCCTTTCCGAGAGGC
TCTTCGCCAACCTGTGGGGGAGGCTTGAGGGGGAGGAGAGGCTCCTTTGGCTTTA
15 CCGGGAGGTGGAGAGGCCCCCTTCCGCTGTCCTGGCCCACATGGAGGCCACGGG
GGTGC GCCTGGACGTGGCCTATCTCAGGGCCTTGTCCCTGGAGGTGGCCGAGGA
GATCGCCCGCCTCGAGGCCGGGTCTTCCGCCTGGCCGGCCACCCCTTCAACCTCA
ACTCCCGGGACCAGCTGGAAAGGGTCCTCTTTGACGAGCTAGGGCTTCCCGCCAT
CGGCAAGACGGAGAAGACCGGCAAGCGCTCCACCAGCGCCGCGCTCCTGGAGGC
20 CCTCCGCGAGGCCCAACCCATCGTGGAGAAGATCCTGCAGTACCGGGAGCTCAC
CAAGCTGAAGAGCACCTACATTGACCCCTTGCCGGACCTCATCCACCCAGGACG
GGCCGCCTCCACACCCGCTTCAACCAGACGGCCACGGCCACGGGCAGGCTAAGT
AGCTCCGATCCCAACCTCCAGAACATCCCCGTCCGCACCCCGCTTGGGCAGAGGA
TCCGCCGGGCCTTCATCGCCGAGGAGGGGTGGCTATTGGTGGCCCTGGACTATAG
25 CCAGATAGAGCTCAGGGTGCTGGCCACCTCTCCGGCGACGAGAACCTGATCCG
GGTCTTCCAGGAGGGGGCGGGACATCCACACGGAGACCGCCAGCTGGATGTTCCG
CGTCCCCCGGGAGGCCGTGGACCCCTGATGCGCCGGGCGGCCAAGACCATCAA
CTTCGGGGTCCTCTACGGCATGTGCGGCCACCGCCTCTCCCAGGAGCTAGCCATC
CCTTACGAGGAGGCCAGGCCTTCATTGAGCGCTACTTTCAGAGCTTCCCCAAGG
30 TGCGGGCCTGGATTGAGAAGACCCTGGAGGAGGGCAGGAGGCGGGGGTACGTG
GAGACCCTCTTCGGCCGCCGCGCTACGTGCCAGACCTAGAGGCCCGGGTGAAG
AGCGTGCGGGAGGCGGCCGAGCGCATGGCCTTCAACATGCCCGTCCAGGGCACC
GCCGCCGACCTCATGAAGCTGGCTATGGTGAAGCTTCCCCAGGCTGGAGGAA

090705301

ATGGGGGCCAGGATGCTCCTTCAGGTCCACGACGAGCTGGTCCTCGAGGCCCA
 AAAGAGAGGGCGGAGGCCGTGGCCCGGCTGGCCAAGGAGGTCATGGAGGGGGT
 GTATCCCCTGGCCGTGCCCCTGGAGGTGGAGGTGGGGATAGGGGAGGACTGGCT
 CTCCGCCAAGGAGGGCATTGATGGCCGCGGCGGAGGCGGGCATCATCATCA
 5 TCATTAA

SEQ ID NO:10 The amino acid sequence of the Sac7d-ΔTaq fusion protein

MITNSTVKVKFKYKGEEKEVDTSKIKKVWRVGMVSFTYDDNGKTGRGAVSEKDA
 PKELLDMLARAEREKKGGGVTSKALEEAPWPPPEGAFVGVLSRKEPMWADLLAL
 10 AAARGGRVHRAPEPYKALRDLKEARGLLAKDLSVLALREGLGLPPGDDPMLLAYLL
 DPSNTTPEGVARRYGGEWTEEAGERAAALSERLFANLWGRLEGEERLLWLYREVERP
 LSAVLAHMEATGVRLDVAYLRALSLEVAEEIARLEAEVFRLAGHPFNLNSRDQLERV
 LFDELGLPAIGKTEKTGKRSTSAAVLEALREAHPIVEKILQYRELTKLKSTYIDPLPDLI
 HPRTGRLHTRFNQTATATGRLSSSDPNLQNIPTVPLGQRIRRAFIAEEGWLLVALDY
 15 SQIELRVLAHLSGDENLIRVFQEGRDIHTETASWMFGVPREAVDPLMRRAAKTINFG
 VLYGMSAHRLSQELAIPYEEAQAFIERYFQSFPKVRAWIEKTLEEGRRRGYVETLFRG
 RRYVPDLEARVKSVREAAERMAFNMPVQGTAAADLMKLMVCLFPRLEEMGARMML
 LQVHDELVLAPKERAEEAVARLAKEVMIEGVYPLAVPLEVEVGIGEDWLSAKEGIDG
 RGGGGHHHHHHH
 20

SEQ ID NO:11 The DNA sequence encoding the PL-ΔTaq fusion protein

ATGATTACGAATTCGAAGAAAAAGAAAAAGAAAAAGCGTAAGAAACGCAAAAA
 GAAAAAGAAAGGCGGCGGTGTCACTAGTGGCGCAACCGTAAAGTTCAAGTACAA
 AGGCGAAGAAAAAGAGGTAGACATCTCCAAGATCAAGAAAGTATGGCGTGTGG
 25 GCAAGATGATCTCCTTCACCTACGACGAGGGCGGTGGCAAGACCGGCCGTGGTG
 CGGTAAGCGAAAAGGACGCGCCGAAGGAGCTGCTGCAGATGCTGGAGAAGCAG
 AAAAAGGGCGGCGGTGTCAACAGTCCCAAGGCCCTGGAGGAGGCCCCCTGGCCC
 CCGCCGGAAGGGGCCTTCGTGGGCTTTGTGCTTTCCCGCAAGGAGCCCATGTGGG
 CCGATCTTCTGGCCCTGGCCGCCGCGCAGGGGGGGCCGGGTCCACCGGGCCCCCG
 30 AGCCTTATAAAGCCCTCAGGGACCTGAAGGAGGCGCGGGGGCTTCTCGCCAAAG
 ACCTGAGCGTTCTGGCCCTGAGGGAAGGCCTTGGCCTCCCGCCCGGCGACGACCC
 CATGCTCCTCGCCTACCTCCTGGACCCTTCCAACACCACCCCGAGGGGGTGGCC
 CGGCGCTACGGCGGGGAGTGGACGGAGGAGGCGGGGGAGCGGGCCGCCCTTTCC

bioRxiv preprint doi: <https://doi.org/10.1101/050235>; this version posted September 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GAGAGGCTCTTCGCCAACCTGTGGGGGAGGCTTGAGGGGGAGGAGAGGCTCCTT
 TGGCTTTACCGGGAGGTGGAGAGGCCCCTTTCCGCTGTCCTGGCCCACATGGAGG
 CCACGGGGGTGCGCCTGGACGTGGCCTATCTCAGGGCCTTGTCCTGGAGGTGGC
 CGAGGAGATCGCCCGCCTCGAGGCCGAGGTCTTCCGCCTGGCCGGCCACCCCTTC
 5 AACCTCAACTCCCGGGACCAGCTGGAAAGGGTCCTCTTTGACGAGCTAGGGCTTC
 CCGCCATCGGCAAGACGGAGAAGACCGGCAAGCGCTCCACCAGCGCCGCCGTCC
 TGGAGGCCCTCCGCGAGGCCACCCCATCGTGGAGAAGATCCTGCAGTACCGGG
 AGCTACCAAGCTGAAGAGCACCTACATTGACCCCTTGCCGGACCTCATCCACCC
 CAGGACGGGGCCGCCTCCACACCCGCTTCAACCAGACGGCCACGGCCACGGGCAG
 10 GCTAAGTAGCTCCGATCCCAACCTCCAGAACATCCCCGTCCGCACCCCGCTTGGG
 CAGAGGATCCGCCGGGCCTTCATCGCCGAGGAGGGGTGGCTATTGGTGGCCCTG
 GACTATAGCCAGATAGAGCTCAGGGTGCTGGCCACCTCTCCGGCGACGAGAAC
 CTGATCCGGGTCTTCCAGGAGGGGGCGGGACATCCACACGGAGACCGCCAGCTGG
 ATGTTTCGGCGTCCCCCGGGAGGCCGTGGACCCCTGATGCGCCGGGGCGGCCAAG
 15 ACCATCAACTTCGGGGTCCTCTACGGCATGTCGGCCACCGCCTCTCCCAGGAGC
 TAGCCATCCCTTACGAGGAGGCCCAGGCCTTCATTGAGCGCTACTTTCAGAGCTT
 CCCCAGGTGCGGGCCTGGATTGAGAAGACCCTGGAGGAGGGCAGGAGGCGGG
 GGTACGTGGAGACCCTCTTCGGCCGCCGCGCTACGTGCCAGACCTAGAGGCC
 GGGTGAAGAGCGTGCGGGAGGCGGCCGAGCGCATGGCCTTCAACATGCCCGTCC
 20 AGGGCACCGCCGCCGACCTCATGAAGCTGGCTATGGTGAAGCTCTTCCCCAGGCT
 GGAGGAAATGGGGGCCAGGATGCTCCTTCAGGTCCACGACGAGCTGGTCCTCGA
 GGCCCCAAAAGAGAGGGCGGAGGCCGTGGCCCGGCTGGCCAAGGAGGTCATGG
 AGGGGGTGTATCCCCTGGCCGTGCCCTGGAGGTGGAGGTGGGGATAGGGGAGG
 ACTGGCTCTCCGCCAAGGAGGGCATTGATGGCCGCGGCGGAGGCGGGCATCATC
 25 ATCATCATCATTA

SEQ ID NO:12 The amino acid sequence of PL-ΔTaq fusion protein

MITNSKKKKKKRKKRKKKKKGGGVTS GATVKFKYKGEEKEVDISKIKKVWRVGK
 MISFTYDEGGGKTGRGAVSEKDAPKELLQMLEKQKKG GVTSPKALEEAPWPPPEG
 30 AFVGFVLSRKEPMWADLLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLSVL
 ALREGLGLPPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEEAGERAALSERLFAN
 LWGRLEGEERLLWLYREVERPLSAVLAHMEATGVRLDVAYLRALSLEVAEEIARLE
 AEVFRLAGHPFNLNSRDQLERVLFDELGLPAIGKTEKTGKRSTSAAVLEALREAHPIV

EKILQYRELTKLKSTYIDPLPDLIHPRTGRLHTRFNQTATATGRLSSSDPNLQNIPVRTP
 LGQRIRRAFIAEEGWLLVALDYSQIELRVLAHLSGDENLIRVFQEGRDIHTETASWMF
 GVPREAVDPLMRRAAKTINFGVLYGMSAHRLSQELAIPYEEAQAFIERYFQSFPKVR
 AWIEKTLEEGRRRGYVETLFGRRRYVPDLEARVKSVREAAERMAFNMPVQGTAAAD
 5 LMKLAMVKLFPRL EEMGARMLLQVHDEL VLEAPKERAEAVARLAKEVMEGVYPL
 AVPLEVEVGIGEDWLSAKEGIDGRGGGGHHHHHH

SEQ ID NO:13 PRIMER L71F

5'-CCTGCTCTGCCGCTTCACGC-3'

10

SEQ ID NO:14 PRIMER L71R

5'-GCACAGCGGCTGGCTGAGGA-3'

SEQ ID NO:15 PRIMER L18015F

15 5'-TGACGGAGGATAACGCCAGCAG-3'

SEQ ID NO:16 PRIMER L23474R

5'-GAAAGACGA TGGGTCGCTAATACGC-3'

SEQ ID NO:17 PRIMER L18015F

20 5'-TGACGGAGGATAAC GCCAGCAG-3'

SEQ ID NO:18 PRIMER L29930R

5'-GGGGTTGGAGGTCAATGGGTTC-3'

25

SEQ ID NO:19 PRIMER L30350F

5'-CCTGCTCTGCCGCTTCACGC-3'

SEQ ID NO:20 PRIMER L35121R

30 5'-CACATGGTACAGCAAGCCTGGC-3'

SEQ ID NO:21 PRIMER L2089F

5'-CCCGTATCTGCTGGGA TACTGGC-3

09870353 053001

SEQ ID NO: 22 PRIMER L7112R

5'-CAGCGGTGCTGACTGAATCATGG-3'

SEQ ID NO:23 PRIMER L30350F

5 5'-CCTGCCTGCCGCTTCACGC-3'

SEQ ID NO:24 PRIMER L40547R

5'-CCAATACCCGTTTCA TCGCGGC-3'

10 **SEQ ID NO:25 PRIMER H-Amelo-Y**

5'-CCACCTCATCCTGG GCACC-3'

SEQ ID NO:26 PRIMER H-Amelo-YR

5'-GCTTGAGGCCAACCATCAGAGC-3'

SEQ ID NO:27 Human beta-globin primer 536F

5'-GGTTGGCCAATCTACTCCCAGG-3'

SEQ ID NO:28 Human beta-globin primer 536R

5'-GCTCACTCAGTGTGGCAAAG-3'

SEQ ID NO:29 Human beta-globin primer 1408R

5'-GATTAGCAAAAGGGCCTAGCTTGG-3'

098703205001
T00E50" E5E0/860

25